MATH 180A: INTRO TO PROBABILITY
(FOR DATA SCIENCE)

www.math.ucsd.edu/~tkemp/180A

Today: § 1.3 - 1.4

Next: § 2.1 - 2.2

HW.0: double check!

HW.1 due FRIDAY, 10/04

Lab.1 due MONDAY, 10/07

Screencast & video available after each lecture @ podcast.ucsd.edu

Before/After slides now available on course webpage.

Lots of active discussion on Piazza.

Combinatorics

* selecting k objects from among n, with replacement:

$$\#\text{ways} = n^k$$

* selecting k objects from among n, without replacement; order matters:

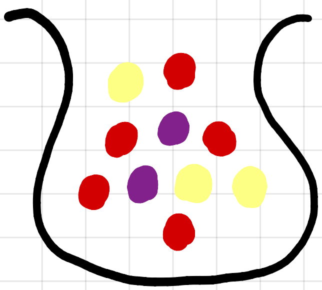
$$\#\text{ways} = n(n-1)(n-2)\cdots(n-k+1) \qquad (k \le n)$$

* selecting k objects from among n, without replacement; order doesn't matter:

$$\#\text{ways} = \binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}$$

$$= \frac{n!}{(n-k)!\,k!} = \binom{n}{n-k}$$

Sampling with Replacement (order doesn't matter)

E.g.  An urn contains 10 balls: $\rightarrow$ $b_1, b_2, b_3\ b_4\ b_5$

2 blue $\qquad b_6\ b_7\ b_8\ b_9\ b_{10}$

3 yellow

5 red

Problem: 3 balls are chosen without replacement.

$$P(2 \text{ yellow}, 1 \text{ red})$$

$$\Omega = \{\{b_1, b_2, b_3\} : b_i \neq b_j \text{ if } i \neq j\} \qquad \#\Omega = \binom{10}{3}$$

↑ order matters

$$A = \{2 \text{ are yellow}, 1 \text{ red}\} \qquad \#A = \binom{5}{1}\cdot\binom{3}{2}$$

$$\therefore\ P(A) = \frac{\binom{5}{1}\cdot\binom{3}{2}}{\binom{10}{3}} = \frac{15}{120} = \frac{1}{8} = 12.5\%$$

1.3

What if $\#\Omega = \infty$?

Then we need a different notion of uniform.

Eg. A random real number is chosen in $[0,1]$.

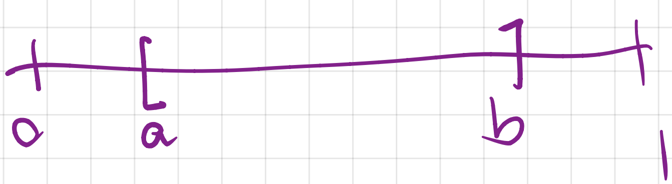
(a) What is the probability it is ≥ 0.7?

(b) What is the probability it is $= \frac{1}{2}$?

$\underbrace{\text{the probability}}_{\text{must define!}}$

$$(\Omega, \mathcal{F}, \mathbb{P})$$

$\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}(\cancel{\mathbb{R}})\ [0,1]$, $\mathbb{P}([a,b]) := b-a.$

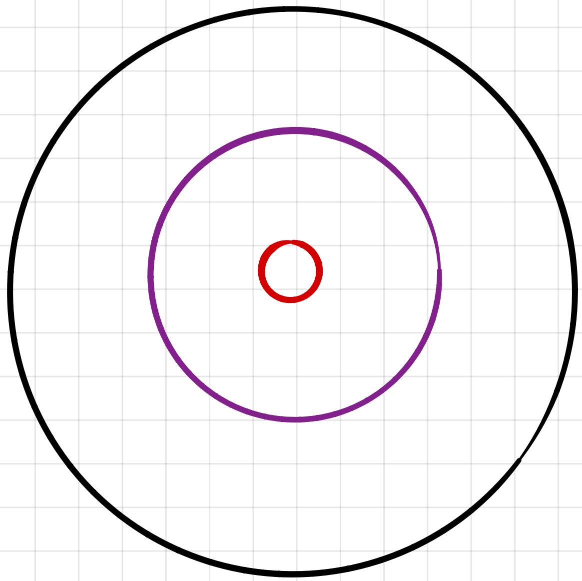


(a) $\mathbb{P}([0.7,1])$
$= 1 - 0.7 = 0.3$

(b) $\mathbb{P}([\frac{1}{2}, \frac{1}{2}])$
$= \frac{1}{2} - \frac{1}{2} = 0.$

$$\mathbb{P}([0,0.3) \cup [\tfrac{1}{2}, 0.96))$$
$$= \mathbb{P}([0,0.3)) + \mathbb{P}([\tfrac{1}{2}, 0.96)) = 0.3 + 0.46$$
$$= 0.76$$

E.g.



An archery target is a disk 50 cm in diameter.

A blue disk in the center is 25 cm in diameter.

A red disk in the center is 5 cm in diameter.

Given that you hit the target (randomly), what are the chances of hitting the blue disk? The red disk?

Ω = target

$\mathcal{F}$ = { subsets that have "area" }

$$\mathbb{P}(A) = \frac{\text{Area}(A)}{\text{Area}(\Omega)}$$

$\mathbb{P}(\text{bullseye}) = 1\%$

1.4

Decompositions

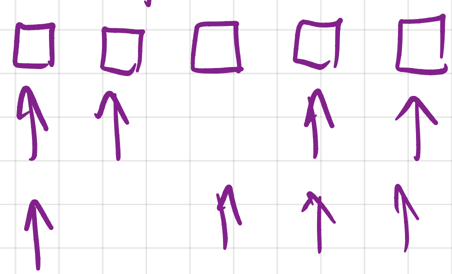
Eg. A fair coin is tossed 5 times. What is the probability that at least 3 tosses come up tails?

$A = \{\text{at least 3 tails}\} = A_3 \cup A_4 \cup A_5$

$A_k := \{\text{exactly } k \text{ tails}\}$

$P(A) = P(A_3) + P(A_4) + P(A_5)$

$P(A_5) = \binom{5}{5}\frac{1}{2^5}$



A_4: # configurations $= \binom{5}{4}$

$P(A_4) = \binom{5}{4}\cdot\frac{1}{2^5}$.

$P(A_3) = \binom{5}{3}\frac{1}{2^5}$

$\therefore P(A) = \frac{1}{2^5}\left(\binom{5}{3}+\binom{5}{4}+\binom{5}{5}\right) = \frac{1}{2^5}(1+5+10) = \frac{16}{32}$

$= 50\%$

E.g. A fair die is rolled 4 times. What is the probability of at least one double?

$A = \{$ some number comes up at least two times $\}$

$A_k = \{$ k comes up at least two times $\}$

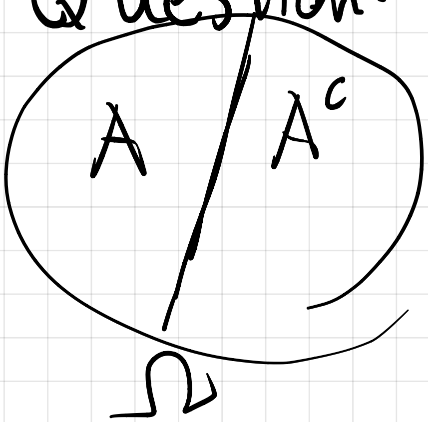
$A = A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 \cup A_6$ not disjoint

$A_k^m = \{$ k comes up exactly m times $\}$ ~~zillions of~~ 30 scenarios.

$A_1 = A_1^2 \cup A_1^3 \cup A_1^4 \cup A_1^5 \cup A_1^6$

$\vdots$

Question: Are all these events disjoint? No!



$$P(A^c) = \frac{6 \cdot 5 \cdot 4 \cdot 3}{6^4} = \frac{5}{18}$$

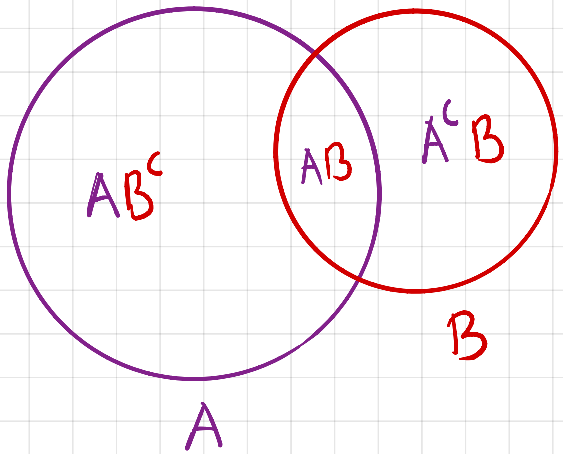
$$1 = P(\Omega) = P(A) + P(A^c)$$

$$\therefore P(A) = 1 - P(A^c) = 1 - \frac{5}{18} = \frac{13}{18}$$

Sometimes, you can't avoid lack of disjointness so easily. You have to take intersections into account.

Notation: $A \cap B = \{\text{all outcomes in both } A \text{ and } B\}$

$$\begin{array}{c} A\cap B \\ \parallel \\ AB \end{array}$$

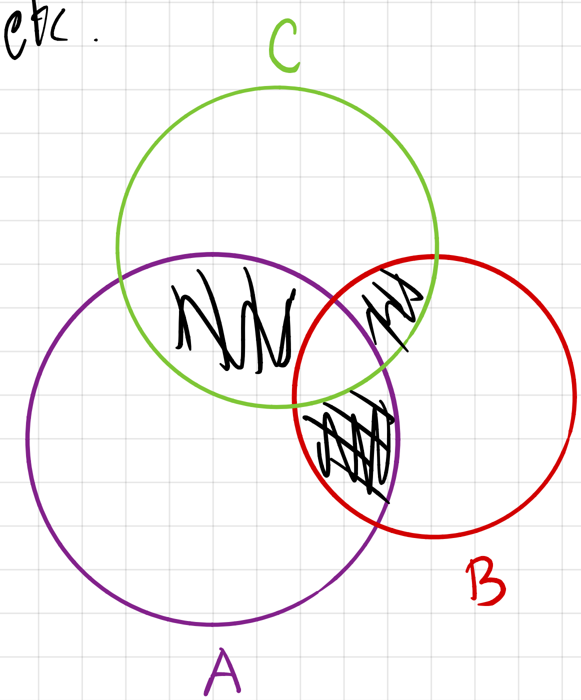


$A \cup B = AB^c \cup AB \cup A^cB \leftarrow$ disjoint

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

Principle of Inclusion / Exclusion

The probability of a union can be computed by adding the probabilities, then subtracting off the intersection(s) overcounted. If you have more sets, you have to keep going and re-add back in pieces that you over-subtracted, etc.



$$\mathbb{P}(A \cup B \cup C)$$
$$= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C)$$
$$- \mathbb{P}(AB) - \mathbb{P}(AC) - \mathbb{P}(BC)$$
$$+ \mathbb{P}(ABC)$$

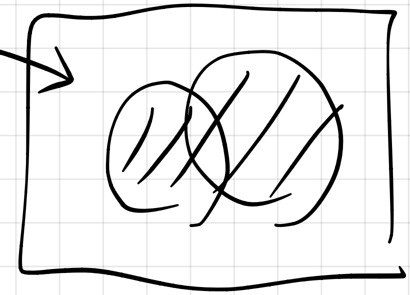
E.g. 20% of the population own cats.
25% of the population own dogs.
5% of the population own both.
What is the probabality that a random person owns neither?

$C \quad D$

$\mathbb{P}(C) = 0.2$

$\mathbb{P}(D) = 0.25$

$\mathbb{P}(CD) = 0.05$



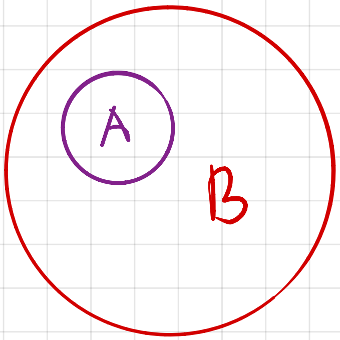
$$
\begin{aligned}
\mathbb{P}(C^cD^c) &= \mathbb{P}((C \cup D)^c) \\
&= 1 - \mathbb{P}(C \cup D) \\
&= 1 - (\mathbb{P}(C) + \mathbb{P}(D) - \mathbb{P}(CD)) \\
&= 1 - (0.2 + 0.25 - 0.05) \\
&= 0.6.
\end{aligned}
$$

Monotonicity

If $A \subseteq B$ then $B = A \cup A^c B$ is a disjoint union

$$\therefore \mathbb{P}(B) = \mathbb{P}(A) + \underbrace{\mathbb{P}(A^c B)}_{\geq 0}$$

$$\geq \mathbb{P}(A)$$



Eg 90% of your friends like the xiao long bao at Din Tai Fung.
80% of your friends like the xiao long bao at Shanghai Saloon.
What is the smallest possible proportion of your friends who like the xiao long bao at both restaurants?